FW

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/751,836	12/29/2000	Philipp Heinz Schmid	M61.12-0323	9264

27366 7590 04/28/2005

MICROSOFT CORPORATION C/O WESTMAN
CHAMPLIN & KELLY, P.A.
SUITE 1600 - INTERNATIONAL CENTRE
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402-3319

EXAMINER	
SKED, MATTHEW J	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 04/28/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/751,836	SCHMID ET AL.	
	Examiner	Art Unit	
	Matthew J Sked	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *26 November 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-39,47 and 50-54* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *47 and 50-54* is/are allowed.

6)☒ Claim(s) *1-3 and 6-39* is/are rejected.

7)☒ Claim(s) *4 and 5* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *26 November 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *11/26/04*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. The replacement drawings were received on 11/26/04. The objection is withdrawn.

2. Applicant's arguments, filed 11/26/04, with respect to the double patenting rejection have been fully considered and are persuasive. The double patenting rejection of claim 1 has been withdrawn.

3. The objections to the specification are withdrawn in view of the amendments filed 11/26/04.

4. The rejection to claim 36 under 35 U.S.C. § 112 has been withdrawn in view of the amendment to the claim.

5. The indicated allowability of claims 1-39 is withdrawn in view of the reference(s) "Sophisticated Speech Processing Middleware on Microprocessor", cited in the previous action. Rejections based on the reference(s) follow.

6. Claims 40-46, 48 and 49 were canceled.

7. Claim 47 has been amended to include an allowable dependent claim and intervening claims hence making the claim allowable for reasons stated in the previous office action.

8. Claim 50 is allowable because it further limits the claim in which it refers.

9. Claims 51-54 remain allowable as presented in the previous office action.

Claim Rejections - 35 USC § 102

10. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

11. Claims 1, 10, 11, 13, 15, 16, 18, 20, 36, 37 and 39 are rejected under 35 U.S.C. 102(a) as being anticipated by Hataoka et al. "Sophisticated Speech Processing Middleware on Microprocessor", cited in the previous action.

Hataoka teaches a middleware layer configured to facilitate communication between a speech-related application and a speech-related engine, comprising:

a speech component having an application-independent interface (multiple applications such as car navigation systems and handheld PC hence making the system application independent, introduction) configured to be coupled to the application and an engine-independent interface (speech recognition and speech synthesis hence making the system engine-independent, introduction) configured to be coupled to the engine and at least one processing component configured to perform speech related services for the application and the engine (speech processing middleware that performs spectral subtraction or prosody mapping processing between the application and the engines, introduction).

12. As per claim 2, Hataoka teaches a marshaling component, configured to access at least one processing component in each process and to marshal information transfer among the processes (middleware performs separate processing for both the speech

recognizer and the synthesizer hence it would inherently transfer the data to the appropriate processors and engines, introduction).

13. As per claim 10, Hataoka teaches a site object having an interface configured to receive result information from the engine (middleware acts as the intermediary between the applications and the engines hence it would inherently be able to receive result information from the engine in order to forward it to the application, introduction).

14. As per claim 11, Hataoka teaches the engine comprises a TTS engine (performs Text-to-Speech synthesis which would inherently need a TTS engine, speech synthesis middleware, page 694) and wherein the processing component comprises a first object having an application interface and an engine interface (middleware consists of processing for both recognition and synthesis hence both would have interfacing to the appropriate engines and be able to transmit the results back to the application, introduction).

15. As per claim 13, Hataoka teaches wherein the application interface exposes a method configured to receive audio device attributes from the application and instantiate a specific audio device based on the audio device attributes received (application switches between arbitrary speech synthesis and fixed sentence speech synthesis, the engines for these specifications would have different attributes, speech synthesis middleware, page 694).

16. As per claim 15, Hataoka teaches the engine interface is configured to call a method exposed by the engine to begin synthesis (middleware performs a natural prosody mapping process during speech synthesis and because there is only one

engine this processing would be exposed when synthesis is chosen, speech synthesis middleware, page 694).

17. As per claim 16, Hataoka teaches the engine comprises a speech recognition engine (speech recognition would inherently have a recognition engine, speech recognition middleware, page 693) and wherein the processing component comprises a first object having an application interface and an engine interface (middleware consists of processing for both recognition and synthesis hence both would have interfacing to the appropriate engines and be able to transmit the results back to the application, introduction).

18. As per claim 18, Hataoka teaches wherein the application interface exposes a method configured to receive audio device attributes from the application and instantiate a specific audio device based on the audio device attributes received (application switches between arbitrary speech synthesis and fixed sentence speech synthesis, the engines for these specifications would have different attributes, speech synthesis middleware, page 694).

19. As per claim 20, Hataoka teaches the application interface exposes a method configured to receive an audio information request from the application and to configure the speech component to retain audio information recognized by the SR engine based on the audio information request (system would inherently buffer or store the inputted voice in order to recognize it, introduction).

20. As per claim 36, Hataoka teaches a site object exposing an engine interface configured to receive information from the speech-related engine (middleware acts as

the intermediary between the applications and the engines hence it would inherently be able to receive result information from the engine in order to forward it to the application, introduction).

21. As per claim 37, Hataoka teaches the site object is configured to receive result information from the SR engine indicative of recognized speech (system has speech recognition capabilities, speech recognition middleware, page 693).

22. As per claim 39, Hataoka teaches a result object configured to obtain the result information from the site object and expose an interface configured to pass the result information to the application (middleware acts as the intermediary between the applications and the engines hence it would inherently be able to forward the result to the application, introduction).

Claim Rejections - 35 USC § 103

23. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

24. Claims 3, 6-9, 12 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hataoka in view of Comerford et al. (U.S. Pat. 6,513,009), cited in the previous office action.

As per claims 3 and 6, Hataoka does not teach a format negotiation component configured to negotiate a data format of data used by the audio device and data used by

the engine wherein the format negotiation component is configured to invoke a format converter to convert the data format of data between the engine and the audio device to a desired format based on the data format used by the audio device and the data format used by the engine.

Comerford teaches a middleware between an application and multiple engines that includes a codec for a speaker that would be used to convert the data from the spoken language engines into a format that could be used for the speaker (Fig. 1B, element 1140).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to include a format converted to convert the format of the data between the engine and the audio device as taught by Comerford because it would allow a digital computer to be used with analog speakers hence allowing commonly used components to be integrated together.

25. As per claims 7 and 8, Hataoka does not teach a lexicon container object configured to contain a plurality of lexicons and to provide a lexicon interface to the engine to represent the plurality of lexicons as a single lexicon to the engine and load the one or more user lexicons as one or more application lexicons.

Comerford teaches a lexicon container object configured to contain a plurality of lexicons and to provide a lexicon interface to the engine to represent the plurality of lexicons as a single lexicon to the engine and load the one or more user lexicons as one or more application lexicons (user interface table has multiple vocabularies that uses an

interpreter to select between them to send to the appropriate targets, col. 6, lines 27-56).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to include a lexicon container object configured to contain a plurality of lexicons and to provide a lexicon interface to the engine to represent the plurality of lexicons as a single lexicon to the engine and load the one or more user lexicons as one or more application lexicons as taught by Comerford because it would expand the systems recognition capabilities hence giving better results.

26. As per claim 9, Hataoka and Comerford do not specifically teach or suggest the lexicon interface is configured to be invoked by the engine to add a lexicon provided by the engine.

However, the Examiner takes Official Notice that choosing a lexicon based upon the type of engine that is being used is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka and Comerford so the lexicon interface is configured to be invoked by the engine to add a lexicon provided by the engine because it would ensure that the current engine is using an appropriate vocabulary hence limiting possible inherent errors in the system.

27. As per claim 12, Hataoka does not specifically teach the application interface exposes a method configured to receive engine attributes from the application and instantiate a specific engine based on the engine attributes received.

Comerford teaches that different applications have different vocabularies hence the determined application would indicate the vocabulary to be used by the speech engine (Fig. 2).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the application interface exposes a method configured to receive engine attributes from the application and instantiate a specific engine based on the engine attributes received as taught by Comerford because it would allow special vocabularies for each application hence making the vocabularies smaller and speeding up recognition.

28. As per claim 22, Hataoka does not teach the engine interface is configured to call the SR engine to set acoustic profile information in the SR engine.

Comerford teaches the engine interface is configured to call the SR engine to set acoustic profile information in the SR engine (during initialization the system incorporates SLI data in the engine which includes acoustic information, col. 7, line 66 to col. 8, line 4 and Fig. 4).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the engine interface is configured to call the SR engine to set acoustic profile information in the SR engine as taught by Comerford because it would allow for speaker-dependent recognition which is more robust.

29. Claims 14, 21, 24 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hataoka.

As per claim 14, Hataoka does not specifically teach a parser to receive input data to be synthesized and parse the input data into text fragments.

However, the Examiner takes Official Notice that parsing text prior to synthesis is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to parse the input data into text fragments because this would allow smaller sounds such as phones to be used to correspond to the text hence giving more natural sounding synthesis.

30. As per claim 21, Hataoka does not teach wherein the application interface exposes a method configured to receive bookmark information from the application identifying a position in an input data stream being recognized and to notify the application when the SR engine reaches the identified position.

However, the Examiner takes Official Notice that notifying a user when a particular position in recognition is reached is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka wherein the application interface exposes a method configured to receive bookmark information from the application identifying a position in an input data stream being recognized and to notify the application when the SR engine reaches the identified position because this would give status feedback to the user hence allowing the user to judge the time remaining in recognition.

31. As per claim 24, Hataoka does not teach the engine interface is configured to call the SR engine to load a language model in the SR engine.

However, the Examiner takes Official Notice that using multiple language models for speech recognition is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to load a language model in the SR engine because it would enable the system to choose a language model better trained in particular environment or for a particular user hence giving better results.

32. As per claim 39, Hataoka does not teach the engine interface on the site object is configured to receive update information from the SR engine indicative of a current position of the SR engine in an audio input stream to be recognized.

However, the Examiner takes Official Notice that notifying a user of recognition status is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka wherein the engine interface on the site object is configured to receive update information from the SR engine indicative of a current position of the SR engine in an audio input stream to be recognized because this would give status feedback to the user hence allowing the user to judge the time remaining in recognition.

33. Claims 17, 19, 23 and 25-35 and are rejected under 35 U.S.C. 103(a) as being unpatentable over Hataoka in view of Baker et al. (U.S. Pat. 6,456,974), cited in the previous office action.

As per claim 17, Hataoka does not teach wherein the application interface exposes a method configured to receive recognition attributes from the application and instantiates a specific speech recognition engine based on the engine attributes received.

Baker teaches a speech recognition system with a middleware for multiple applications that receives recognition attributes from the application and instantiates a specific speech recognition engine based on the engine attributes received (application specifies the grammar, col. 3, lines 14-16).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to receive recognition attributes from the application and instantiate a specific speech recognition engine based on the engine attributes received as taught by Baker because it would allow the system to use a grammar associated with the application hence giving better results.

34. As per claim 19, Hataoka does not teach wherein the application interface exposes a method configured to receive an alternate request from the application and to configure the speech component to retain alternates provided by the SR engine for transmission to the application based on the alternate request.

Baker teaches the application interface exposes a method configured to receive an alternate request from the application and to configure the speech component to retain alternates provided by the SR engine for transmission to the application based on the alternate request (SRResult class gives a list of n-best results and confidence

scores based upon the recognition of the input which is sent back to the application, col. 5, lines 3-6 and Fig. 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka wherein the application interface exposes a method configured to receive an alternate request from the application and to configure the speech component to retain alternates provided by the SR engine for transmission to the application based on the alternate request as taught by Baker because it would give the application alternate results to choose from if the recognized request is incorrect.

35. As per claim 23, Hataoka does not teach the engine interface is configured to call the SR engine to load a grammar in the SR engine.

Baker teaches the engine interface is configured to call the SR engine to load a grammar in the SR engine (col. 3, lines 14-22).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the engine interface is configured to call the SR engine to load a grammar in the SR engine as taught by Baker because it would allow the system to use a grammar associated with the application hence giving better results.

36. As per claims 25 and 27, Hataoka does not teach wherein the application interface exposes a method configured to receive a grammar request from the application and to instantiate a grammar object based on the grammar request to be used by the SR engine.

Baker teaches the application interface exposes a method configured to receive a grammar request from the application and to instantiate a grammar object based on the grammar request to be used by the SR engine (SRGrammar class specifies the grammar from the application which is used in recognition, col. 3, lines 14-22 and col. 4, line 66 to col. 5, line 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the application interface exposes a method configured to receive a grammar request from the application and to instantiate a grammar object based on the grammar request as taught by Baker because it would allow the system to use a grammar associated with the application hence giving better results.

37. As per claim 26, Hataoka does not specifically teach the grammar object includes a word sequence data buffer.

Baker teaches the speech recognizer has access and uses a grammar in recognition (col. 6, lines 13-15). A grammar looks at the words before and after a recognized word in order to aide in recognition. Therefore, a grammar would inherently have a word sequence data buffer.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the grammar object includes a word sequence data buffer as taught by Baker because it would enable the system to evaluate the surrounding vocabulary hence allowing the best recognition result to be obtained.

38. As per claim 28, Hataoka does not teach the grammar includes words, rules and transitions and wherein the grammar object includes an application interface and an engine interface.

Baker teaches the grammar includes words, rules and transitions (grammar is a set of rules defining syntax and vocabulary, col. 3, lines 16-17) and wherein the grammar object includes an application interface and an engine interface (SpeechRecognizer class controls the connection between the recognition engine and application, col. 4, lines 62-66).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka so the grammar includes words, rules and transitions and wherein the grammar object includes an application interface and an engine interface as taught by Baker because it would allow the grammar rules to added and deleted easily.

39. As per claims 29 and 31, Hataoka does not teach wherein the application interface exposes a grammar configuration method configured to receive grammar configuration information from the application and configure the grammar based on the grammar configuration information, wherein the grammar configuration method is configured to receive grammar activation information and enable or disable grammars in the grammar object based on the grammar activation information.

Baker teaches wherein the application interface exposes a grammar configuration method configured to receive grammar configuration information from the application and configure the grammar based on the grammar configuration information,

wherein the grammar configuration method is configured to receive grammar activation information and enable or disable grammars in the grammar object based on the grammar activation information (SRGrammar can activate and deactivate grammars, col. 4, line 66 to col. 5, line 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to receive grammar activation information and enable or disable grammars in the grammar object based on the grammar activation information as taught by Baker because it would allow the system to dynamically alter the recognition grammars hence making the system more adaptable to changes in recognition conditions.

40. As per claims 30 and 32, neither Hataoka nor Baker specifically teach grammar configuration method is configured to receive word change data, rule change (activation/deactivation) data, and transition change data and change words, rules and transitions in the grammar in the grammar object based on the grammar received.

However, Baker teaches reloading an altered grammar dynamically (col. 4, line 66 to col. 5, line 3). Because the grammar includes words, rules and transitions this would suggest changing the words, rules and transitions of the grammar.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to receive word change data, rule change (activation/deactivation) data, and transition change data and change words, rules and transitions in the grammar in the grammar object based on the grammar received as suggested by Baker because it would allow the system to dynamically alter the

recognition grammars hence making the system more adaptable to changes in recognition conditions.

41. As per claims 33-35, Hataoka does not teach the engine interface is configured to call the SR engine to load the grammar in the SR engine, wherein the call updates a configuration of the grammar or activation state in the SR engine.

Baker teaches the engine interface is configured to call the SR engine to load the grammar in the SR engine, wherein the call updates a configuration of the grammar or activation state in the SR engine (API communicates the grammar to use to the SR engine wherein the call includes a data structure that specifies activating, deactivating and altering grammars, col. 3, lines 14-22 and col. 4, line 66 to col. 5, line 6).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Hataoka to call the SR engine to load the grammar in the SR engine as taught by Baker because it would allow the system to dynamically alter the recognition grammars hence making the system more adaptable to changes in recognition conditions.

Allowable Subject Matter

42. Claims 4 and 5 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims. The prior art on record does not specifically teach reconfiguring the data used by an audio device or the speech engine.

Conclusion

43. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J Sked whose telephone number is (571) 272-7627. The examiner can normally be reached on Mon-Fri (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David L Ometz can be reached on (571) 272-7593. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MS
04/25/05

DAVID L. OMETZ
PRIMARY EXAMINER